



JC03 Rec'd PCT/PTO 30 NOV 2001

FORM PTO-1390 (REV. 9-2001)		U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER P/126-213	
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371				U.S. APPLICATION NO. (If known, see 37 CFR 1.5) 09/980275	
INTERNATIONAL APPLICATION NO. PCT/JP99/03492		INTERNATIONAL FILING DATE May 31, 2000		PRIORITY DATE CLAIMED 05.11.1999 & 10.20.99	
TITLE OF INVENTION DEVICE, METHOD, AND PROGRAM FOR ENCODING/DECODING OF SPEECH WITH FUNCTION OF ENCODING SILENT PERIOD					
APPLICANT(S) FOR DO/EO/US Masahiro SERIZAWA and Hironori ITO					
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:					
1. <input checked="" type="checkbox"/> This is a FIRST submission of items concerning a filing under 35 U.S.C. 371. 2. <input type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371. 3. <input type="checkbox"/> This is an express request to begin national examination procedures (35 U.S.C. 371(f)). The submission must include items (5), (6), (9) and (21) indicated below. 4. <input checked="" type="checkbox"/> The US has been elected by the expiration of 19 months from the priority date (Article 31). 5. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2)) a. <input type="checkbox"/> is attached hereto (required only if not communicated by the International Bureau). b. <input checked="" type="checkbox"/> has been communicated by the International Bureau. c. <input type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US). 6. <input checked="" type="checkbox"/> An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)). a. <input checked="" type="checkbox"/> is attached hereto. b. <input type="checkbox"/> has been previously submitted under 35 U.S.C. 154(d)(4). 7. <input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3)) a. <input type="checkbox"/> are attached hereto (required only if not communicated by the International Bureau). b. <input type="checkbox"/> have been communicated by the International Bureau. c. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired. d. <input checked="" type="checkbox"/> have not been made and will not be made. 8. <input type="checkbox"/> An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371 (c)(3)). 9. <input checked="" type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)). - unsigned 10. <input type="checkbox"/> An English language translation of the annexes of the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).					
Items 11 to 20 below concern document(s) or information included:					
11. <input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98. 12. <input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included. 13. <input type="checkbox"/> A FIRST preliminary amendment. 14. <input type="checkbox"/> A SECOND or SUBSEQUENT preliminary amendment. 15. <input type="checkbox"/> A substitute specification. 16. <input type="checkbox"/> A change of power of attorney and/or address letter. 17. <input type="checkbox"/> A computer-readable form of the sequence listing in accordance with 37 CFR 1.96. 18. <input type="checkbox"/> A second copy of the published international application and English language translation. 19. <input type="checkbox"/> A second copy of the English language translation of the international application. 20. <input checked="" type="checkbox"/> Other items or information: 10 sheets of drawings PEFS print form PCT/IB/308					
EXPRESS MAIL CERTIFICATE I hereby certify that this correspondence is being deposited with the United States Postal Service as Express Mail Post Office Addressee (Mail Label EL 924390248 US) in an envelope addressed to: U.S. Patent and Trademark Office, PO Box 2327 Arlington, VA 22202, on November 30, 2001					
Tamika Sumpter Name of Person Mailing correspondence					
 Signature					
November 30, 2001 Date of Signature					

U.S. APPLICATION NO. (if known) 09/1780275		INTERNATIONAL APPLICATION NO. PCT/JP00/03492		ATTORNEY'S DOCKET NUMBER P/126-213	
21. <input checked="" type="checkbox"/> The following fees are submitted: BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)): Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO..... \$1040.00 International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO \$890.00 International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO \$740.00 International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4) \$710.00 International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4) \$100.00 ENTER APPROPRIATE BASIC FEE AMOUNT =				CALCULATIONS PTO USE ONLY	
Surcharge of \$130.00 for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(e)).				\$ 890.00	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE	\$	
Total claims	73 - 20 =	53	x \$18.00	\$ 954.00	
Independent claims	18 - 3 =	15	x \$84.00	\$ 1260.00	
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ \$280.00	\$	
TOTAL OF ABOVE CALCULATIONS =				\$ 3104.00	
<input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.				+	
SUBTOTAL =				\$ 3104.00	
Processing fee of \$130.00 for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(f)).				\$	
TOTAL NATIONAL FEE =				\$ 3104.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property +				\$	
TOTAL FEES ENCLOSED =				\$ 3104.00	
				Amount to be refunded:	\$
				charged:	\$
a. <input checked="" type="checkbox"/> A check in the amount of \$ <u>3104.</u> to cover the above fees is enclosed. Check No. <u>7530</u>					
b. <input type="checkbox"/> Please charge my Deposit Account No. _____ in the amount of \$ _____ to cover the above fees. A duplicate copy of this sheet is enclosed.					
c. <input checked="" type="checkbox"/> The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. <u>15-0700</u> . A duplicate copy of this sheet is enclosed.					
d. <input type="checkbox"/> Fees are to be charged to a credit card. WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.					
NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137 (a) or (b)) must be filed and granted to restore the application to pending status.					
SEND ALL CORRESPONDENCE TO: OSTROLENK, FABER, GERB & SOFFEN, LLP 1180 Avenue of the Americas New York, NY 10036-8403 Tel: (212) 382 0700					
				 SIGNATURE	
				Edward A. Meilman NAME	
				24,735 REGISTRATION NUMBER	

13 Rec'd PCT/PTO 20 FEB 2002

09/980275

#2

DICKSTEIN SHAPIRO MORIN & OSHINSKY LLP

1177 Avenue of the Americas • 41st Floor • New York, New York 10036-2714

Tel (212) 835-1400 • Fax (212) 997-9880

Writer's Direct Dial. (212) 896-5471

E-Mail Address: MeilmanE@dsmo.com

February 20, 2002

BY FACSIMILE No. 703-305-3230

Ms. Donna Green
U.S. Patent and Trademark Office
PO Box 2327
Arlington, VA 22202

Re: DSMO Ref. No. G0126.0213 (Formerly P/126-213)
National Phase in U.S.A. for PCT/JP00/03492
Inventors: Masahiro SERIZAWA, et al
DEVICE METHOD, AND PROGRAM FOR ENCODING/
DECODING OF SPEECH WITH

Dear Ms. Green:

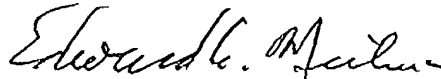
Confirming our telephone conversation, the designation of the International Application Number on the first page of Form PTO-1390 is not correct while the designation of the International Application Number on the second page of that form is correct. Would you please change the designation of the International Application Number on the first page that form to PCT/JP00/03492.

In addition, my correspondence address, telephone number and fax number have changed. The new contact information is:

Dickstein Shapiro Morin & Oshinsky LLP
1177 Avenue of the Americas - 41st Floor
New York, New York 10036
Phone: (212) 896-5470
Fax: (212) 997-9880

Would you please correct the PTO records to reflect this new information.

Sincerely yours,



Edward A. Meilman

EAM/jst/mgs

2101 L Street NW • Washington, DC 20037-1526

Tel (202) 785-9700 • Fax (202) 887-0689

www.legalinnovators.com

10/19/01

09980275, 051602
09/980275

SPECIFICATION

JC03 Rec'd PCT/PTO 30 NOV 2001

TITLE OF THE INVENTION

DEVICE, METHOD, AND PROGRAM FOR ENCODING / DECODING
OF SPEECH WITH FUNCTION OF ENCODING SILENT PERIOD

TECHNICAL FIELD

The invention relates to a device for encoding/decoding of digital information such as a speech signal, in particular, to a technique for encoding/decoding of a voice-less period.

BACKGROUND ART

Conventionally, some devices are proposed to reduce an average bit rate of transmission of a speech signal in a voice-less period (a period with no voice), by encoding a speech signal at lower bit rates than that used to encode a speech signal in a period with a voice. For example, the technique is disclosed in a document 1 (IEEE Communication Magazine, pages 64 - 73, Sep. 1997).

The conventional encoding device determines whether the input signal includes a voice or not, for each frame with a predetermined size, e.g. 10 milliseconds, and if the signal in the frame includes a voice, the signal is encoded and decoded in a general speech coding method.

On the other hand, the input signal includes no voice, the conventional coding device discontinuously encodes feature parameters of the input speech signal and transmits the encoded parameters to a decoding device. Herein, the decoding device smoothes the feature parameters discontinuously received, and decodes a speech signal by using the smoothed parameters.

A method of determining whether the speech signal is voice-less or not for each frame, is also disclosed in the document 1. In the method, a root means square value (hereinafter, referred to as "RMS") computed from an input speech signal for each frame, an RMS corresponding to a low frequency region, the number of zero crossing, and filter coefficients representing spectral envelope characteristics are used.

The determination is done by comparing these values in each frame with the predetermined thresholds.

A method of encoding a speech signal in a period with voice is, for example, disclosed as CELP method (Code Excited Linear Prediction Coding method) in a document 2 (ITU-T recommendation G.729, July, 1995).

The CELP method is disclosed in a document 3 (Code-Excited Linear Prediction: High Quality Speech at Very Low Bit Rates (IEEE Proc. ICASSP-85, pp. 937 - 940, 1985)).

In an encoding process of a conventional coding device, first, speech signal is inputted frame by frame and is processed with linear predictive analysis to obtain linear predictive (LP) coefficients representing spectral envelope characteristics of a speech, and an excitation signal for driving an LP synthesis filter corresponding to the spectral envelope characteristics is derived to be encoded.

Further, in an encoding process of the excitation signal, each frame is divided into subframes and encoding of the excitation signal is performed for each subframe. Herein, the excitation signal is composed of a pitch element representing a pitch period of the input signal, a residual element, and gains of these elements. The pitch element is denoted as an adaptive codevector which is stored in a codebook, which is referred to as "adaptive codebook", and includes the past

excitation signal. The residual element is denoted as a multipulse signal composed of a plurality of pulses.

Also, in a decoding process, to decode a speech signal, an excitation signal derived by decoding the pitch element and the residual element is fed into a synthesis filter composed of decoded filter coefficients.

In a method of encoding a speech signal in a voice-less period, as described in the document 1, first, an RMS and filter coefficients calculated from the speech are encoded at a coding device. Then, at a decoding device, a multipulse signal and a random signal are generated so that a root mean square of a sum of them is equal to the decoded RMS, and the sum of them is fed to a synthesis filter composed using the decoded filter coefficients to decode a speech signal in a voice-less period.

In a voice-less period, the feature parameters are transmitted only in frames that characteristics of the signal changes, otherwise nothing is transmitted. However, information showing whether the feature parameters is transmitted or not is sent in another way.

When the feature parameters are not transmitted, the output speech signal is decoded by repeatedly using the past transmitted feature parameters. Smoothed RMS is used for decoding not to cause a discontinuity of a waveform of the decoded speech signal.

Fig. 8 shows a block diagram representing a structure of a conventional encoding device. Referring to Fig. 8, the encoding device includes a voice part coding circuit 12, a voice-less part coding circuit 14, a signal determining circuit 16, a switching circuit 18, and a bit sequence generating circuit 20.

A speech signal is inputted frame by frame, for example, in 10 milliseconds unit by an input terminal 10. The signal determining circuit

16 determines whether the speech signal from the input terminal 10 is a period with voice or a voice-less period for each frame, and passes the determining result (VAD determination sign) to the switching circuit 18 and a bit sequence generating circuit 20.

The voice part coding circuit 12 encodes the speech signal from the input terminal 10 for each frame, and passes the encoded signal to the switching circuit 18.

The voice-less part coding circuit 14 encodes the speech signal from the input terminal 10 for each frame, and passes the encoded signal to the switching circuit 18. Further, the voice-less part coding circuit 14 sends determination information (DTX determination sign) indicating whether the encoded signal is transmitted in the voice-less period, to the bit sequence generating circuit 20.

The switching circuit 18 operates based on the VAD determination sign received from the signal determining circuit 16. When the circuit 18 receives the sign indicating a voice period, the encoded signal passed from the voice part coding circuit 12 is sent to the bit sequence generating circuit 20. On the other hand, when the circuit 18 receives the sign indicating a voice-less period, the encoded signal passed from the voice-less part coding circuit 14 is sent to the bit sequence generating circuit 20.

The bit sequence generating circuit 20 multiplexes the VAD determination sign from the signal determining circuit 16, the DTX determination sign from the voice-less part coding circuit 10, and encoded signal from the switching circuit 18, to generate bit sequence and outputs the bit sequence from an output terminal 22.

Fig. 9 shows a block diagram for explaining a conventional decoding device.

Referring to Fig. 9, the decoding device includes a bit sequence decomposing circuit 26, a switching circuit 28, a voice part decoding circuit 30, and a voice-less part decoding circuit 34.

The bit sequence decomposing circuit 26 decomposes a bit sequence inputted from an input terminal 24 into the VAD determination sign, the DTX determination sign, and the encoded signal. And then, the circuit 26 sends the VAD determination sign and the encoded signal to the switching circuit 28, and sends the DTX determination sign to the voice-less part decoding circuit 34.

The switching circuit 28 operates based on the VAD determination sign received from the bit sequence decomposing circuit 26. When the circuit 28 receives the sign indicating a voice period, the encoded signal passed from the bit sequence decomposing circuit 26 is sent to the voice part decoding circuit 30. On the other hand, when the circuit 28 receives the sign indicating voice-less period, the encoded signal passed from the bit sequence decomposing circuit 26 is sent to the voice-less part decoding circuit 34.

The voice part decoding circuit 30 decodes the encoded signal passed from the switching circuit 28 and outputs the decoded signal from an output terminal 32.

The voice-less part decoding circuit 34 decodes the encoded signal passed from the switching circuit 28 by using the DTX determination sign from the bit sequence decomposing circuit 26, and outputs the decoded signal from an output terminal 32.

Fig. 10 shows a block diagram representing a voice-less part decoding circuit 34 of a conventional decoding device. Referring to Fig. 10, the voice-less part decoding circuit 34 includes a parameter decoding circuit 54, a random circuit 56, a pulse circuit 53, a pitch circuit 58, a mixing circuit 61, a smoothing circuit 66, and a synthesis circuit 68.

The parameter decoding circuit 54 decodes filter coefficients and an RMS from the encoded signal inputted from an input terminal 52, and sends the filter coefficients and the RMS to the synthesis circuit 68 and the smoothing circuit 66, respectively.

The smoothing circuit 66 receives the RMS from the parameter decoding circuit 54, and smoothes the RMS. And then the circuit 66 passes the smoothed RMS to the mixing circuit 61. However, if it is found that the encoded signal is not transmitted through the DTX determination sign from an input terminal 50, the circuit 66 calculates the smoothed RMS by smoothing the RMS values of the past frames.

Herein, a smoothed RMS $P(n)$ which is used in the n -th frame in a voice-less period is calculated by using the following equation (1) with the RMS $p(n)$ received in the n -th frame. However, when no encoded signal is transmitted, the RMS of the previous frame is used in the equation (1) instead of $p(n)$.

$$P(n) = (1 - \alpha) \cdot P(n-1) + \alpha \cdot p(n) \dots (1)$$

Herein, α is a smoothing factor for determining a degree of smoothing, in the above-mentioned document 1, a fixed value 0.125 is set. Further, $P(-1)$ is equal to zero.

The random circuit 56 generates a random signal and passes the random signal to the mixing circuit 61. The pulse circuit 53 generates a multipulse signal composing of a plurality of pulses, each of which has a location and an amplitude determined based on each random number, and passes the multipulse signal to the mixing circuit 61.

The pitch circuit 58 generates a pitch signal $q(i)$ composed of the above-mentioned adaptive codevector, and passes it to the mixing circuit 61. Since a pitch period used to define the adaptive codevector is not transmitted, a random number is used instead.

The mixing circuit 61 computes an excitation signal $x(i)$ to be fed into a synthesis filter by performing the linear sum of the random signal $r(i)$ from the random circuit 56, the multipulse signal $p(i)$ from the pulse circuit 53, and the pitch signal $q(i)$ from the pitch circuit 58, and the result of the computation is sent to the synthesis circuit 68.

A method can be used of computing coupling coefficients of the linear sum as described in the document 1.

In the method, first, a coupling coefficient of the pitch signal G_q is selected from a limited range of values according to a random number.

Next, using the G_q , a coupling coefficient of the multipulse signal G_p is calculated so that the RMS derived from the linear sum of the pitch signal and the multipulse signal is equal to the smoothed RMS.

Using thus calculated G_q and G_p , the linear sum of the pitch signal and the multipulse signal $e(i)$ is calculated according to the following equation (2).

$$e(i) = G_q \cdot q(i) + G_p \cdot p(i) \dots (2)$$

Furthermore, a coupling coefficient of the linear sum of $e(i)$ and the random signal $r(i)$, $G_r(i)$ and γ , is computed so that the RMS derived from the linear sum of the $e(i)$ and $r(i)$ is equal to the smoothed RMS. Herein, as a coupling coefficient of the random signal, a fixed value, $\gamma = 0.6$ is used.

Therefore, the excitation signal to be fed into the synthesis filter, $x(i)$, is computed according to the following equation (3).

$$x(i) = G_r \cdot [G_q \cdot q(i) + G_p \cdot p(i)] + \gamma \cdot r(i) \dots (3)$$

The synthesis circuit 68 decodes the encoded signal by feeding the excitation signal passed from the mixing circuit 61 to a synthesis filter composed of the filter coefficients passed from the parameter decoding circuit 54. Then, the circuit 68 outputs the decoded speech

signal from an output terminal 70.

However, the above-mentioned conventional device includes the following problems.

The first problem is that there may be a case where filter coefficients used to decode a speech signal in a voice-less period changes discontinuously at a decoding device, and therefore, degradation of a quality of decoded signal occurs.

That reason is because discontinuously transmitted filter coefficients are used as they are.

The second problem is that a decoding process in the beginning period (for example, several hundreds of milliseconds) in a voice-less period may be influenced by a voice period right before the voice-less period, and consequently an amplitude of the decoded signal is increased over the actual amplitude or degradation of speech quality of the decoded signal occurs, for example, due to existence of echoed sound.

That reason is because a smoothing process of the RMS is always performed in a voice-less period to prevent decoded (reproduced) signals in the voice-less period from being discontinuous.

The third problem is that decoded signal in a voice-less period is remarkably different from a background noise of input speech signal in hearing the decoded signal, and as a result, discontinuous auditory impression is given between the background noise included in the voice-less period and a background noise in a voice period.

That reason is because a fixed value is used as a ratio of a pulse element and a pitch element to a random element, in generating an excitation signal to be fed into the synthesis filter in a voice-less period.

Therefore, the invention is considering the problems. It is a main object of the invention to encode a speech signal in a voice-less period

in a high performance, and to provide a device which realizes a high coding quality even if an average transmission bit rate is decreased to encode a speech signal in a voice-less period.

It is another object of the invention to provide a decoding device which can reduce a degradation of the speech quality due to discontinuity of the filter coefficients in decoding a speech signal in a voice-less period.

DISCLOSURE OF THE INVENTION

According to a first aspect of the invention to realize the objects, a speech decoding device is provided, which changes a decoding operation of a speech signal according to whether the speech signal is in a voice period or in a voice-less period in each frame, and which selects feature parameters representing spectral envelope characteristics of the speech signal to be decoded from the feature parameters, smoothes the selected feature parameters in a time direction, and decodes the speech signal by using the smoothed feature parameters.

According to a second aspect of the invention, a speech decoding device is provided which changes a decoding operation of a speech signal according to whether the speech signal is in a voice period or in a voice-less period in each frame, and which decodes the speech signal by using a value, which is obtained by smoothing, in a time direction, at least one of the feature parameters according to an elapsed time from a time point when a transition occurs from the voice period to the voice-less period.

According to a third aspect of the invention, a speech decoding device is provided which changes a decoding operation of a speech signal according to whether the speech signal is in a voice period or in a

voice-less period in each frame, and which decodes the voice signal by using a value, which is obtained from at least one of the received feature parameters as it is in a certain time period immediately after changing from the voice period to the voice-less period, and obtained by smoothing at least one of the feature parameters in a time period after the certain time period.

According to a fourth aspect of the invention, a speech decoding device is provided which changes a decoding operation of a speech signal according to whether the speech signal is in a voice period or in a voice-less period in each frame, and which decodes the speech signal by using a value, which is obtained by smoothing at least one of the feature parameters according to the feature parameters.

According to a fifth aspect of the invention, a speech decoding device is provided which changes a decoding operation of a speech signal according to whether the speech signal is in a voice period or in a voice-less period in each frame, and which decodes the speech signal by using a value, which is obtained by smoothing, in a time direction, at least one of the feature parameters according to at least one of the feature parameters and an elapse time from when a transition is made from a voice period to a voice-less period.

According to a fifth aspect of the invention, a speech decoding device is provided which changes a decoding operation of a speech signal according to whether the speech signal is in a voice period or in a voice-less period in each frame, and which decodes the speech signal by using a value, which is obtained from at least one of the feature parameters as it is when the feature parameter satisfies a predetermined condition, and obtained by smoothing, in a time direction, at least one of the feature parameters after the condition is not satisfied.

According to a sixth aspect of the invention, a speech decoding device is provided which changes a decoding operation of a speech signal according to whether the speech signal is in a voice period or in a voice-less period in each frame, and which decodes the speech signal by using a value which is obtained by smoothing, in a time direction, at least one of the feature parameters according to an elapsed time from when a transition is made from a voice period to a voice-less period.

According to a seventh aspect of the invention, a speech decoding device is provided which changes a decoding operation of a speech signal according to whether the speech signal is in a voice period or in a voice-less period in each frame, and which decodes the speech signal by using a value, which is obtained from at least one of the feature parameters as it is when the feature parameter satisfies a predetermined condition and immediately after a transition is made from a voice period to a voice-less period, otherwise, obtained by smoothing, in a time direction, at least one of the feature parameters.

According to an eighth aspect of the invention, a speech decoding device is provided, which changes a decoding operation of a speech signal according to whether the speech signal is in a voice period or in a voice-less period in each frame, and which generates the a speech signal in a part of a voice-less period by feeding an excitation signal composed of plural types of signals, and determines coefficients used to perform a sum operation of the plural types of signals according to at least one of the received feature parameters.

According to a ninth aspect of the invention, a speech decoding device is provided, which changes a decoding operation of the speech signal according to whether the speech signal is in a voice period or in a voice-less period in each frame, and which generates a speech signal in a voice-less period by feeding an excitation signal composed of plural

types of signals, and determines, in a part of the period, a coefficient used to perform a sum operation of the plural types of signals according to at least one of the feature parameters smoothed in a time direction.

According to a tenth aspect of the invention, in the speech decoding device of the above the first aspect to the ninth aspect, the feature parameter includes at least one of a quantity representing spectral envelope of the signal to be decoded and a quantity representing power of the signals to be decoded.

According to an eleventh aspect of the invention, a coding device which determines whether the speech signal is in a voice period or in a voice-less period in each frame, and encodes a feature parameter of the speech signal is incorporated with the voice decoding device of the first aspect to the tenth aspect.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows a diagram of a structure of a voice-less part decoding circuit according to a first embodiment of the invention.

Fig. 2 shows a diagram of a structure of a decoding device according to a second embodiment of the invention.

Fig. 3 shows a diagram of a structure of a voice-less part decoding circuit according to a second embodiment of the invention.

Fig. 4 shows a diagram of a structure of a decoding device according to a third embodiment of the invention.

Fig. 5 shows a diagram of a structure of a voice-less part decoding circuit according to a third embodiment of the invention.

Fig. 6 shows a diagram of a structure of a decoding device according to a fourth embodiment of the invention.

Fig. 7 shows a diagram of a structure of a voice-less part decoding circuit according to a fourth embodiment of the invention.

Fig. 8 shows a diagram of a structure of a coding device according to a conventional device and the invention.

Fig. 9 shows a diagram of a structure of a conventional decoding device.

Fig. 10 shows a diagram of a structure of a voice-less part decoding circuit of a conventional decoding device.

BEST MODE FOR EMBODYING THE INVENTION

Description is made about embodiments of the invention. A speech decoding device according to a first embodiment of the invention includes a switching device (shown in Fig. 9 (28)), a smoothing device (shown in Fig. 1 (64)), and a group of decoding devices (shown in Fig. 1 (56, 53, 58, 61, and 68)).

The switching device switches the method of decoding the signal by using the feature parameters of the encoded signal to be decoded, according to determination information representing whether the encoded signal is in a voice period or in a voice-less period for each frame. The smoothing device smoothes the feature parameters representing spectral envelope characteristics of the encoded signal. The group of decoding devices decodes the encoded signal by using the smoothed feature parameters.

A speech decoding device according to a second embodiment of the invention includes a switching device (shown in Fig. 2 (28)), a group of smoothing devices (shown in Fig. 2 (36) and Fig. 3 (49 and 51)), and a group of decoding devices (shown in Fig. 3 (56, 53, 58, 61, and 68)).

The switching device switches the method of decoding the signal by using the feature parameters of encoded signal to be decoded, according to determination information representing whether the encoded signal is in a voice period or in a voice-less period for each

frame. The group of smoothing devices smooths at least one parameter in the feature parameters, based on the parameters and an elapsed time from a time point when a voice period is changed to a voice-less period. The group of decoding devices decodes the encoded signals by using the smoothed feature parameters.

A speech decoding device according to a third embodiment of the invention includes a switching device (shown in Fig. 2 (28)), a group of smoothed value generating devices (shown in Fig. 2 (36) and Fig. 3 (49 and 51)), and a group of decoding devices (shown in Fig. 3 (56, 53, 58, 61, and 68)).

The switching device switches methods of decoding the signal by using feature parameters of encoded signals to be decoded, according to determination information representing whether the encoded signal is in a voice period or in a voice-less period for each frame. The group of smoothed value generating devices set the original value of at least one of transmitted feature parameters as a smoothed value immediately after transition from a voice period to a voice-less period and when a feature parameter satisfies predetermined conditions, and thereafter, generate a smoothed value by smoothing at least one of the feature parameters. The group of decoding devices decodes the encoded signals by using the smoothed parameters.

A speech decoding device according to a fourth embodiment of the invention includes a switching device (shown in Fig. 4 (28)), a group of signal generating devices (shown in Fig. 5 (56, 53, 58, 60, and 68)), and a coefficient determining device (shown in Fig. 5 (38)).

The switching device switches the method of decoding the signal by using the feature parameters of encoded signals to be decoded, according to determination information representing whether the encoded signal is in a voice period or in a voice-less period for each

frame. The group of signal generating devices generates a decoded signal of a voice-less period by feeding an excitation signal composed of plural types of signals into a synthesis filter. The coefficient determining device determines coefficients used to mix plural types of signals in the voice-less period according to at least one of the received feature parameters.

A speech decoding device according to a fifth embodiment of the invention includes a switching device (shown in Fig. 6 (28)), a group of signal generating devices (shown in Fig. 7 (56, 53, 58, 62, and 68)), a group of parameter calculating devices (shown in Fig. 7 (49 and 51)), and a coefficient determining device (shown in Fig. 6 (38)).

The switching device switches methods of decoding signals by using feature parameters of encoded signals to be decoded, according to determination information representing whether the encoded signal is in a voice period or in a voice-less period for each frame. The group of signal generating devices generates a signal of a voice-less period by feeding an excitation signal composed of plural types of signals into a synthesis filter. The group of parameter calculating devices calculates a smoothed parameter by smoothing the received feature parameters. The coefficient determining device determines coefficients used to mix plural types of signals in the voice-less period according to at least one of the calculated feature parameters.

In a speech decoding device according to a sixth embodiment of the invention, the feature parameters include at least one of a value representing the spectral envelope of the signals to be decoded and a value representing a power of the signals.

A preferred embodiment of a encoding/decoding device according to the invention includes a encoding device (shown in Fig. 8) which determines whether the input signal is in a voice period or in a

voice-less period for each frame and encodes feature parameters of the input signal, and a speech decoding device according to one of the devices shown in the first embodiment to the sixth embodiment.

Description is made about an operation and a principle of an embodiment of the invention.

According to the invention, the speech decoding device smoothes a discontinuously transmitted filter coefficients with the RMS, and uses the coefficients about a synthesis filter, in decoding a speech signal in a voice-less period. Thereby, a discontinuous change of the filter coefficients can be prevented which is caused due to the discontinuous transmission of the filter coefficients, and as a result, a voice quality of the decoded signal can be improved.

In the speech decoding device, when the filter coefficients and the RMS which are smoothed in a voice-less period are currently used, the filter coefficients and the RMSs of the past frames influence the currently used filter coefficients and the RMS because of the smoothing process.

Since the signal in the beginning of the voice-less period includes characteristics of a voice period immediately before the voice-less period, the signal in the voice-less period is decoded by using the feature parameters including the characteristics of the voice period. Consequently, an amplitude of a waveform of the decoded signal become larger than an actual amplitude of the input speech signal, or degradation of the decoded speech signal, such as an existence of echo in the decoded signal, may occur.

To prevent them, when a predetermined time elapses or a certain number of frames are received from a time point of the transition from a voice period to a voice-less period, for example, a smoothing factor is set not to perform smoothing process when a value of the RMS

representing an amplitude of the decoded speech is still larger than a predetermined value. Thereby, in the beginning of the voice-less period, an effect from the voice period immediately before the voice-less period, due to smoothing of the feature parameter can be reduced.

There may be the auditory difference between a background noise included in the signal decoded in a voice part decoding circuit and the signal decoded in a voice-less part decoding circuit, in a case where background noises are included in the input signal. This reason is that the voice-less part decoding circuit computes an excitation signal to be fed into a synthesis filter, on only condition that the RMS of the signal becomes equal to a smoothed value of the transmitted RMS.

In the invention, it is capable of reducing degradation of the decoded speech quality due to the auditory difference, by determining how to compute the excitation signal considering characteristics of the input signal. To consider the characteristics, for example, a random noise signal is mainly used when the smoothed RMS is small, on the other hand, a pulse signal or a pitch signal is mainly used when the smoothed RMS is large or when the spectrum computed from the filter coefficients are not flat.

Description is made in more detail about embodiments of the invention with reference to the drawings. A basic structure of an encoding device used in the embodiments is similar to the structure of the coding device shown in Fig. 8. Also, a basic structure of the decoding device is similar to the structure of the decoding device shown in Fig. 9.

Fig. 1 shows a block diagram of a structure of a voice-less part decoding circuit in a decoding device according to the first embodiment of the invention. Referring to Fig. 1, the voice-less part decoding circuit of the first embodiment is different from the voice-less part decoding

circuit 34 shown in Fig. 10 in that the former voice-less part decoding circuit further includes a smoothing circuit 64. In the following description, it is mainly explained about the difference between the device according to the invention and the conventional device, therefore, explanation about common parts will be omitted.

A parameter decoding circuit 54 determines the filter coefficients and the RMS by using a sequence of signals received from an input terminal 52, and passes the determined filter coefficient and the determined RMS to the smoothing circuit 64 and the other smoothing circuit 66, respectively.

The smoothing circuit 64 smoothes the filter coefficients received from the parameter decoding circuit 54 and passes the smoothed filter coefficients to the synthesis circuit 68. However, the smoothing circuit 64 performs smoothing process by using the filter coefficients of the past frames when the DTX determination sign received from an input terminal 50 indicates that the feature parameters are received.

Smoothed filter coefficients $F(n, i)$, ($i = 1, \dots, M$) used for the n -th frame from the beginning of each voice-less period, is calculated by using an equation (4) with the filter coefficients $f(n, i)$ ($i = 1, \dots, M$) entered in the n -th frame. Also, in a frame where nothing is transmitted, the filter coefficients sent immediately before the frame are used to calculate instead of $f(n, i)$.

$$F(n, i) = (1 - \beta)F(n-1, i) + \beta f(n, i) \dots (4)$$

Herein, β is a smoothing factor to determine a degree of smoothing. Also, $F(-1, i)$, ($i = 1, \dots, M$) is equal to 0.

M is an order of the synthesis filter. The synthesis circuit 68 decodes the signal by feeding an excitation signal received from the mixing circuit 61 into the synthesis filter composed of the filter coefficients received from the smoothing circuit 64, and outputs the

decoded signal to an output terminal 70.

Fig. 2 shows a diagram representing a structure of the decoding device according to the second embodiment of the invention. The embodiment differs from the conventional decoding device shown in Fig. 9 in that a structure of a voice-less part decoding circuit 35 of the embodiment is different from that of the conventional decoding device, and the embodiment includes a smoothing control circuit 36.

Hereinafter, description is mainly made about the difference between the decoding device according to the second embodiment and the conventional decoding device, and explanation about parts each of which is the same as the corresponding part of the conventional decoding device may be omitted for the sake of convenience.

A bit sequence decomposing circuit 26 decomposes a bit sequence supplied from an input terminal 24 into a VAD determination sign, a DTX determination sign, and a sequence of the encoded signal, and passes the VAD determination sign to a smoothing control circuit 36 and a switching circuit 28, passes the sequence of the signal to the switching circuit 28, and passes the DTX determination sign to a voice-less part decoding circuit 35.

The switching circuit 28 passes the sequence of the signal passed from the bit sequence decomposing circuit 26 to a voice part decoding circuit 30 when the VAD determination sign from the bit sequence decomposing circuit 26 indicates that the input signal is in a voice period, or passes the sequence of the signal to a voice-less part decoding circuit 35 when it indicates that input signal is in a voice-less period.

The smoothing control circuit 36 passes smoothing factors $\alpha(n)$ and $\beta(n)$ determined based on a change of the VAD determination sign from the bit sequence decomposing circuit 26, to the voice-less part

decoding circuit 35. Herein, n represents a frame number, counted from the beginning, of frames in each voice-less period.

For example, when the VAD determination sign indicates that the input signal is in a voice-less period, an effect of a part in a voice period immediately before the voice-less period on the beginning part in the voice-less period can be reduced by setting each of values of the smoothing factors $\alpha(n)$ and $\beta(n)$ to 1 in the first specified frames or for a specified period in the voice-less period. Further, by setting each of values of the smoothing factors $\alpha(n)$ and $\beta(n)$ to 1 while a similarly transmitted parameter such as the filter coefficients or the RMS satisfies a specified condition, an effect of a part in a voice period immediately before the voice-less period on the beginning part in the voice-less period can be reduced.

For example, the specified condition is that the RMS is more than a threshold value or that both the RMS and the RMS of the first subframe in the voice-less period are less than a threshold value, for detecting that the RMS is under the influence of the part, in a voice period, immediately before the voice-less period. Also, the specified condition may be that a distance (for example, square distance) between the filter coefficients and a predetermined filter coefficients is less than a predetermined threshold value for detecting that the filter coefficients are similar to a smoothed spectrum in a voice period.

Further, when a voice period immediately before a first voice-less period does not include a certain number of frames or is shorter than a certain length of period, a smoothed value in the last frame of a second voice-less period immediately before the voice period can be used as an initial value $P(-1)$, $F(-1, i)$, ($i = 1, \dots, M$) for calculating smoothed values of the filter coefficients and the RMS, since it is considered that the characteristics of the input signal in the second

voice-less period is similar to the characteristics of the input signal in the first voice-less period.

The voice-less part decoding circuit 35 decodes the signal in a voice-less period by using the smoothing factors $\alpha(n)$ and $\beta(n)$, the DTX determination sign received from the bit sequence decomposing circuit 26, and the sequence of the signal received from the switching circuit 28, and outputs the decoded signal to an output terminal 32.

Fig. 3 shows a diagram representing a structure of the voice-less part decoding circuit 35 according to the second embodiment of the invention. The voice-less part decoding circuit 35 is different from the voice-part decoding circuit of the first embodiment of the invention in a structure of a smoothing circuit 49 and a smoothing circuit 51.

A parameter decoding circuit 54 determines the filter coefficients and the RMS based on a sequence of the encoded signal entered from an input terminal 52, and passes the filter coefficients to the smoothing circuit 49 and passes the RMS to the smoothing circuit 51.

The smoothing circuit 49 smoothes the filter coefficients supplied from the parameter decoding circuit 54 by using a smoothing factor $\beta(n)$ entered from an input terminal 65, and passes the smoothed filter coefficients to a synthesis circuit 68. However, when the DTX determination sign received from an input terminal 50 indicates that the encoded signal is not transmitted the filter coefficients of the previous frame is repeatedly used.

The smoothed filter coefficients used in the n-th frame from the beginning of each voice-less period, $F(n, i)$, ($i = 1, \dots, M$) can be calculated by using the following equation (5) which is similar to the above equation (4), with the filter coefficients entered in the n-th frame $f(n, i)$.

$$F(n, i) = (1 - \beta(n)) \cdot F(n-1, i) + \beta(n) \cdot f(n, i) \dots (5)$$

Herein, a value of $\beta(n)$ is changed according to the number of frames which have already received in each voice-less period, and takes about 1 when a few frames are received, so as to remove an effect from the past frames. For example, it can be set as follows.

$\beta(1)=\beta(2)=1.0, \beta(3)=\beta(4)=\dots=\beta(L)=0.7$. Herein, L is the number of frames in each voice-less period.

The smoothing circuit 51 smoothes the RMS sent from the parameter decoding circuit 54 and passes the smoothed RMS to a mixing circuit 61. However, when the DTX determination sign sent from an input terminal 50 indicates that the encoded signal is not transmitted, a smoothing process is performed by using the RMS recently received. The smoothed RMS $P(n)$, which is used in the n -th frame from the beginning of each voice-less period, is calculated by using the following equation (6) which is similar to the equation (1), with the RMS $p(n)$ entered in the n -th frame.

$$P(n) = (1 - \alpha(n)) \cdot P(n-1) + \alpha(n) \cdot p(n) \dots (6)$$

Herein, similarly to $\beta(n)$, $\alpha(n)$ is changed according to the number of frames which have already received in each voice-less period, and takes about 1 when a few frames are received, so as to remove an effect from the past frames. For example, it can be set as follows.

$\alpha(1)=\alpha(2)=1.0, \alpha(3)=\alpha(4)=\dots=\alpha(L)=0.7$. Herein, L is the number of frames in each voice-less period.

Also, one of the processes of the smoothing circuits 49 and 51 can be performed. In this case, the filter coefficients or the RMS sent from the parameter decoding circuit 54 are or is directly sent to the synthesis circuit 68 or a mixing circuit 61.

In the mixing circuit 61, calculates an excitation signal $x(i)$ to be fed into a synthesis filter by performing the linear sum about a random

signal $r(i)$ sent from a random circuit 56, a pulse signal $p(i)$ sent from a pulse circuit 53, and a pitch signal $q(i)$ sent from a pitch circuit 58 with a smoothed RMS sent from the smoothing circuit 51, and passes the calculated signal to the synthesis circuit 68.

The synthesis circuit 68 decodes the speech signal by feeding the excitation signal sent from the mixing circuit 61 into the synthesis filter composed of the filter coefficients sent from the smoothing circuit 49, and outputs the decoded speech signal from an output terminal 70.

Fig. 4 shows a diagram representing a structure of a decoding device according to the third embodiment of the invention. The embodiment differs from the conventional decoding device in a voice-less part examining circuit 38 and a voice-less part decoding circuit 37.

A bit sequence decomposing circuit 26 decomposes a bit sequence supplied from an input terminal 24 into a VAD determination sign, a DTX determination sign, and a sequence of signals, and passes the VAD determination sign and the sequence of signals to a switching circuit 28, and passes the DTX determination sign to a voice-less part decoding circuit 37.

The switching circuit 28 passes the signal passed from the bit sequence decomposing circuit 26 to a voice part decoding circuit 30 when the VAD determination sign from the bit sequence decomposing circuit 26 indicates that the input signal is in a voice period, or passes the sequence of signals to a voice-less part decoding circuit 37 when it indicates that the input signal is in a voice-less period.

The voice-less part examining circuit 38 determines a set up parameter to adjust coupling coefficients of the linear sum used at the mixing circuit 62 shown in Fig. 5 by using the filter coefficients and the RMS sent from the voice-less part decoding circuit 37, and passes the parameters to the voice-less part decoding circuit 37. Description will

be made later with a process in the mixing circuit 62 about calculation of the set up parameters.

Fig. 5 shows a diagram representing a structure of the voice-less part decoding circuit 37 according to the third embodiment of the invention. The voice-less part decoding circuit 37 is different from the voice-part decoding circuit 35 of the first embodiment of the invention in a mixing circuit 62 and an output destination of a parameter decoding circuit 54. Hereinafter, description is made mainly about the difference, and description about the common part is omitted.

A parameter decoding circuit 54 determines the filter coefficients and the RMS based on a sequence of signals entered from an input terminal 52, and passes the filter coefficients to the smoothing circuit 64 and an output terminal 23, and passes the RMS to the smoothing circuit 66 and an output terminal 25.

The smoothing circuit 66 smoothes the RMS passed from the parameter decoding circuit 54 and passes the smoothed RMS to a mixing circuit 62. When the DTX determination sign sent from an input terminal 50 indicates that the encoded signal is not transmitted, the RMS, which is transmitted immediately before the current frame, is used to smooth. Further, it can be controlled not to update the smoothed RMS by setting smoothing factors $\alpha(n)$ and $\beta(n)$ to zero.

A random circuit 56 generates a random number and passes the random number to the mixing circuit 62.

A pulse circuit 53 generates a pulse signal composed of a pulse having a location and an amplitude generated base on the random number, and passes the pulse signal to the mixing circuit 62.

The mixing circuit 62 calculates coupling coefficients of the above-mentioned linear sum by using the set up parameter received from an input terminal 60 and the smoothed RMS received from the

smoothing circuit 66.

Also, the circuit 62 calculates a linear sum signal of the random signal from the random circuit 56, the pulse signal from the pulse circuit 53, and the pitch signal from the pitch circuit 58 by using the coupling coefficients, and passes the linear sum signal to the synthesis circuit 68.

The synthesis circuit 68 decodes input signal by feeding an excitation signal sent from the mixing circuit 62 into a filter composed of the filter coefficients sent from the smoothing circuit 64, and outputs the decoded signal from an output terminal 70.

Next, description is made about the voice-less part examining circuit 38 and the mixing circuit 62.

The voice-less part examining circuit 38 determines the characteristics of a background noise in a voice-less part, and changes a calculation method of the coupling coefficients of the pitch signal, the pulse signal, and the random signal in the mixing circuit, according to the determined characteristics. As set up parameters to be changed, there are an order to decide the coupling coefficients or a coupling coefficient γ .

The voice-less part examining circuit 38 uses information, for example, the RMS and the filter coefficients to determine the characteristics of the background in the voice-less part.

According to a method of controlling the set up parameters based on the above the illustrated information, when the RMS is less than a predetermined threshold value and thereby it is presumed that there is no background noise, or when it is presumed that the input signal is a white noise since an inclination of spectrum of the input signal calculated from the filter coefficients is flat, a contribution rate of the random signal is expanded. It means that a value of γ is reduced with keeping the order of calculation of the coupling coefficients.

Also, the set up parameters of the voice-less period can be included in a sequence of signals and transmitted with the signals.

Fig. 6 shows a diagram representing a structure of a decoding device according to the fourth embodiment of the invention. The embodiment differs from the second embodiment of the invention in a voice-less part examining circuit 38 and a voice-less part decoding circuit 39.

A bit sequence decomposing circuit 26 decomposes a bit sequence supplied from an input terminal 24 into a VAD determination sign, a DTX determination sign, and a sequence of signals, and passes the VAD determination sign to a smoothing control circuit 36 and a switching circuit 28, passes the sequence of signals to the switching circuit 28, and passes the DTX determination sign to a voice-less part decoding circuit 39.

The switching circuit 28 passes the sequence of signals passed from the bit sequence decomposing circuit 26 to a voice part decoding circuit 30 when the VAD determination sign from the bit sequence decomposing circuit 26 indicates that the encoded signal is in a voice period, or passes the sequence of signals to a voice-less part decoding circuit 39 when it indicates that input signal is in a voice-less period.

The smoothing control circuit 36 passes the smoothing factors α (n) and β (n) which are determined according to a change of the VAD determination sign sent from the bit sequence decomposing circuit 26 to the voice-less part decoding circuit 39.

The voice-less part examining circuit 38 determines a set up parameter to adjust coupling coefficients of the linear sum used at the mixing circuit 62 shown in Fig. 7 by using a smoothed RMS sent from the voice-less part decoding circuit 39, and passes the parameters to the voice-less part decoding circuit 39.

The voice-less part detecting circuit 39 can perform a set up parameter determining process by replacing RMS with smoothed RMS in above-mentioned process of the voice-less part examining circuit 38.

The voice-less part detecting circuit 39 decodes an input signal in a voice-less period, by using the DTX determination sign from the bit sequence decomposing circuit 26, the encoded signal from the switching circuit 28, the smoothing factors $\alpha(n)$ and $\beta(n)$ from the smoothing control circuit 36, and the set up parameters from the voice-less part examining circuit 38, and outputs the decoded signal from an output terminal 32.

Also, smoothed RMS calculated by a smoothing circuit 51 shown in Fig. 7 and smoothed filter coefficients calculated by a smoothing circuit 49 are passed to the voice-less part examining circuit 38.

Fig. 7 shows a diagram representing a structure of the voice-less part decoding circuit 39 according to the fourth embodiment of the invention. The voice-less part decoding circuit 39 is different from the voice-part decoding circuit of the second embodiment of the invention in that in the fourth embodiment, an output from a smoothing circuit 51 is supplied to an output terminal 69 and a smoothing circuit 49 is supplied to an output terminal 63.

In each of the above described embodiments of the invention, a pitch signal, a pulse signal, and a random signal is used to compute an excitation signal of a synthesis filter, but any of them can be omitted.

A decoding device according to the invention and a coding device described in a background section of the specification can be applied to a radio terminal or a radio base station thereby, a radio voice communication system using a speech signal compressing technique can be easily established. Further, a voice terminal can be easily constructed by storing a program to perform the above described

decoding method of the invention into a storage medium such as a floppy disk and by loading the program into a personal computer to which a loudspeaker is connected.

As described above, according to the invention, the following effects are obtained.

A first effect of the invention is that speech quality degradation due to discontinuous change of the filter coefficients used in decoding the signal in a voice-less period can be prevented in the decoding device of the invention.

This reason is that the discontinuously transmitted filter coefficient is smoothed and used in the invention.

A second effect of the invention is that a speech quality degradation due to influence of a voice period immediately before a voice-less period on the beginning of the voice-less period can be reduced in the decoding device of the invention.

This reason is that a smoothing factor is adjusted not to smooth the feature parameters in the beginning of a voice-less period.

A third effect of the invention is that auditory discontinuity caused by a transition between a voice period and a voice-less period can be reduced in the decoding device of the invention.

This reason is that when an excitation signal of a reproduction filter is generated in a voice-less period, ratio of a random element to a pulse element and a pitch element is changed according to a nature of input signals.

CLAIMS

1. A speech decoding device which decodes a speech signal by using received feature parameters according to whether the speech signal is in a voice period or in a voice-less period, the device comprising:

a decoding unit which decodes the speech signal in at least a part of the voice-less period by using a smoothed value obtained by smoothing the feature parameter representing spectral envelope characteristics.

2. A speech decoding device which decodes a speech signal by using received feature parameters according to whether the speech signal to be decoded is in a voice period or in a voice-less period, the device comprising:

a voice-less part decoding unit which changes, according to an elapsed time from a time point when a transition occurs from the voice period to the voice-less period, a coefficient used to smooth at least one of the feature parameters, and decodes the speech signal in the voice-less period by smoothing at least one of the feature parameters with the changed coefficient.

3. The speech decoding device of claim 2, wherein the voice-less part decoding unit decodes the speech signal by using at least one of the received feature parameters as it is in a time period immediately after changing from the voice period to the voice-less period, and decodes the speech signal by using at least one smoothed feature parameter selected from the feature parameters after the time period.

4. A speech decoding device which decodes speech signal by using received feature parameters according to whether the speech

signal to be decoded is in a voice period or in a voice-less period, the device comprising:

a voice-less part decoding unit which changes a value of a coefficient used to smooth at least one of the feature parameters according to the feature parameters, and decodes the speech signal in the voice-less period by smoothing at least one of the feature parameters with the changed value of the coefficient.

5. The speech decoding device of claim 4, wherein the voice-less part decoding unit decodes the speech signal by using at least one of the received feature parameters as it is while a feature parameter satisfies a predetermined condition, and decodes the speech signal by using at least one smoothed feature parameter selected from the feature parameters after the condition is not satisfied.

6. A speech decoding device which decodes speech signal by using received feature parameters according to whether the speech signal to be decoded is in a voice period or in a voice-less period, the device comprising:

a voice-less part decoding unit which changes a value of a coefficient used to smooth at least one of the feature parameters according to information representing whether the feature parameters are transmitted or not, and decodes the speech signal in the voice-less period by smoothing at least one of the feature parameters with the changed value of the coefficient.

7. The speech decoding device of claim 2, wherein the voice-less part decoding unit changes, according to an elapsed time from a time point when a transition occurs from the voice period to the voice-less period and to the feature parameters, a value of a coefficient used to smooth at least one of the feature parameters, and decodes the speech signal in the voice-less period by smoothing at least one of the

feature parameters with the changed value of the coefficient.

8. The speech decoding device of claim 3, wherein when the decoding process of the speech signal in the voice-less period is performed after the time period, the decoding process is performed by using at least one smoothed feature parameter selected from the feature parameters, according to an elapsed time from a time point when a transition occurs from the voice period to the voice-less period and to at least one of the feature parameters.

9. The speech decoding device of claim 5, wherein when the decoding process of the speech signal in the voice-less period is performed after the time period, the decoding process is performed by using at least one smoothed feature parameter selected from the feature parameters, according to an elapsed time from a time point when a transition occurs from the voice period to the voice-less period and to at least one of the feature parameters.

10. The speech decoding device of claim 2, wherein the voice-less part decoding unit decodes the speech signal by using at least one of the received feature parameters as it is, in a first time period immediately after changing from the voice period to the voice-less period and in a second time period while the feature parameter satisfies a predetermined condition, and decodes the speech signal by using at least one smoothed feature parameter selected from the feature parameters after the first time period or the second time period is past.

11. The speech decoding device of claim 2, wherein the voice-less part decoding unit changes a value of a coefficient used to smooth at least one of the feature parameters according to information representing whether the feature parameters are transmitted or not, and decodes the speech signal in the voice-less period by smoothing at least one of the feature parameters with the changed value of the

coefficient.

12. The speech decoding device of claim 4, wherein the voice-less part decoding unit changes a value of a coefficient used to smooth at least one of the feature parameters according to Information representing whether the feature parameters are transmitted or not, and decodes the speech signal in the voice-less period by smoothing at least one of the feature parameters with the changed value of the coefficient.

13. The speech decoding device of claim 6, wherein the voice-less part decoding unit receives information representing whether the feature parameters are sent at a sending location.

14. The speech decoding device of claim 11, wherein the voice-less part decoding unit receives information representing whether the feature parameters are sent at a sending location.

15. The speech decoding device of claim 12, wherein the voice-less part decoding unit receives information representing whether the feature parameters are sent at a sending location.

16. The speech decoding device of claim 1, wherein when a length of a voice period immediately before a first voice-less period is shorter than a predetermined length, a value of a feature parameter which is finally transmitted in a second voice-less period immediately before the voice period is used as an initial value of smoothing.

17. The speech decoding device of claim 2, wherein when a length of a voice period immediately before a first voice-less period is shorter than a predetermined length, a value of a feature parameter which is finally transmitted in a second voice-less period immediately before the voice period is used as an initial value of smoothing.

18. The speech decoding device of claim 4, wherein when a length of a voice period immediately before a first voice-less period is

shorter than a predetermined length, a value of a feature parameter which is finally transmitted in a second voice-less period immediately before the voice period is used as an initial value of smoothing.

19. The speech decoding device of claim 6, wherein when a length of a voice period immediately before a first voice-less period is shorter than a predetermined length, a value of a feature parameter which is finally transmitted in a second voice-less period immediately before the voice period is used as an initial value of smoothing.

20. A speech decoding device which decodes speech signal by using received feature parameters according to whether the speech signal is in a voice period or in a voice-less period, the device comprising:

a voice-less part decoding unit which generates signals in the voice-less period by feeding an excitation signal composed of plural types of signals to a synthesis filter in the voice-less period, wherein the voice-less part decoding unit comprises a weighting coefficient determining unit which determines a weighting coefficient used in a weighted sum operation of the plurality of types of signals in the voice-less period according to at least one feature parameter, and the excitation signal generated by using the weighting coefficient is fed to the synthesis filter.

21. A speech decoding device which decodes speech signal by using received feature parameters according to whether the speech signal is in a voice period or in a voice-less period, the device comprising:

a voice-less part decoding unit which generates signals in the voice-less period by feeding an excitation signal composed of plural types of signals to a synthesis filter in the voice-less period, wherein the voice-less part decoding unit comprises a weighting coefficient

determining unit which determines a weighting coefficient used in a weighted sum operation of the plurality of types of signals in the voiceless period according to at least one smoothed feature parameter obtained by smoothing a feature parameter selected from the received feature parameters in a time direction, and the excitation signal generated by using the weighting coefficient is fed to the synthesis filter.

22. The speech decoding device of claim 1, wherein the feature parameters includes at least one of a quantity representing spectral envelope of the signals to be decoded and a quantity representing power of the signals to be decoded.

23. The speech decoding device of claim 2, wherein the feature parameters includes at least one of a quantity representing spectral envelope of the signals to be decoded and a quantity representing power of the signals to be decoded.

24. The speech decoding device of claim 4, wherein the feature parameters includes at least one of a quantity representing spectral envelope of the signals to be decoded and a quantity representing power of the signals to be decoded.

25. The speech decoding device of claim 6, wherein the feature parameters includes at least one of a quantity representing spectral envelope of the signals to be decoded and a quantity representing power of the signals to be decoded.

26. The speech decoding device of claim 20, wherein the feature parameters includes at least one of a quantity representing spectral envelope of the signals to be decoded and a quantity representing power of the signals to be decoded.

27. The speech decoding device of claim 21, wherein the feature parameters includes at least one of a quantity representing spectral envelope of the signals to be decoded and a quantity

representing power of the signals to be decoded.

28. The speech decoding device of claim 1 being included in a speech coding/decoding device with a coding device which determines whether the input signal is in a voice period or in a voice-less period for each frame and encodes the feature parameters of the input signals to output.

29. The speech decoding device of claim 2 being included in a speech coding/decoding device with a coding device which determines whether the input signal is in a voice period or in a voice-less period for each frame and encodes the feature parameters of the input signals to output

30. The speech decoding device of claim 4 being included in a speech coding/decoding device with a coding device which determines whether the input signal is in a voice period or in a voice-less period for each frame and encodes the feature parameters of the input signals to output.

31. The speech decoding device of claim 6 being included in a speech coding/decoding device with a coding device which determines whether the input signal is in a voice period or in a voice-less period for each frame and encodes the feature parameters of the input signals to output.

32. The speech decoding device of claim 20 being included in a speech encoding/decoding device with a coding device which determines whether the input signal is in a voice period or in a voice-less period for each frame and encodes the feature parameters of the input signals to output.

33. The speech decoding device of claim 21 being included in a speech coding/decoding device with a coding device which determines whether the input signal is in a voice period or in a voice-less period for

each frame and encodes the feature parameters of the input signals to output.

34. A method of decoding speech signal by changing a decoding operation corresponding to received feature parameters according to whether the speech signal is in a voice period or in a voice-less period, the method comprising the steps of:

smoothing the feature parameter representing spectral envelope characteristics of the speech signal to be decoded in at least a part of the voice-less period; and

decoding the speech signal in the voice-less period by using the smoothed feature parameter.

35. A method of decoding speech signal by changing a decoding operation corresponding to received feature parameters according to whether the speech signal is in a voice period or in a voice-less period, the method comprising the steps of:

smoothing at least one of the feature parameters according to an elapsed time from a time point when a transition occurs from the voice period to the voice-less period; and

decoding the speech signal in the voice-less period by using the smoothed feature parameter.

36. The method of claim 35, wherein the smoothing step further comprises the steps of:

(a) providing at least one of the received feature parameters as it is as the smoothed feature parameter in a certain time period immediately after changing from the voice period to the voice-less period; and

(b) smoothing at least one of the received feature parameters in a time period other than the certain time period.

37. A method of decoding speech signal by changing a decoding operation corresponding to received feature parameters according to whether the speech signal is in a voice period or in a voice-less period, the method comprising the steps of:

smoothing at least one of the feature parameters according to the received feature parameters; and

decoding the speech signal in the voice-less period by using the smoothed feature parameter.

38. The method of claim 37, wherein the smoothing step further comprises the steps of:

(a) providing at least one of the received feature parameters as it is as the smoothed feature parameter when a feature parameter satisfies a predetermined condition; and

(b) smoothing at least one of the received feature parameters after the condition is not satisfied.

39. A method of decoding speech signal by changing a decoding operation corresponding to received feature parameters according to whether the speech signal is in a voice period or in a voice-less period, the method comprising the steps of:

smoothing at least one of the feature parameters according to information representing whether the feature parameters are transmitted or not; and

decoding the speech signal in the voice-less period by using the smoothed feature parameter.

40. The method of claim 35, wherein the smoothing step smoothes at least one of the received feature parameters according to an elapsed time from when a transition occurs from the voice period to the voice-less period, and the feature parameters.

41. The method of claim 35, wherein the smoothing step smoothes at least one of the received feature parameters according to an elapsed time from when a transition occurs from the voice period to the voice-less period, and at least one of the feature parameters, after at least one of the feature parameters is used as it is.

42. The method of claim 37, wherein the smoothing step smoothes at least one of the received feature parameters according to an elapsed time from when a transition occurs from the voice period to the voice-less period, and at least one of the feature parameters, after at least one of the feature parameters is used as it is.

43. The method of claim 35, wherein the smoothing step further comprises the steps of:

(a) providing at least one of the received feature parameters as it is, in a first time period immediately after changing from the voice period to the voice-less period and in a second time period while the feature parameter satisfies a predetermined condition; and

(b) smoothing at least one of the received feature parameters in a time direction after the first time period or the second time period is past.

44. The method of claim 35, wherein the smoothing step changes a coefficient used to smooth at least one of the received feature parameters according to information representing whether the feature parameters are transmitted or not.

45. The method of claim 37, wherein the smoothing step changes a coefficient used to smooth at least one of the received feature parameters according to information representing whether the feature parameters are transmitted or not.

46. The method of claim 39 further comprising the step of receiving the information representing whether the feature parameters

are transmitted or not.

47. The method of claim 44 further comprising the step of receiving the information representing whether the feature parameters are transmitted or not.

48. The method of claim 45 further comprising the step of receiving the information representing whether the feature parameters are transmitted or not.

49. A method of decoding speech signal based on received feature parameters by changing a decoding operation according to whether the speech signal is in a voice period or in a voice-less period, at least a part of decoding process in the method in the voice-less period comprising the steps of:

determining a weighting coefficient used to generate an excitation signal of the voice-less period by performing a weighted sum operation of plural types of signals based on at least one of the received feature parameters; and

generating the excitation signal based on the weighting coefficient, and generating speech signal in the voice-less period by feeding the excitation signal to a synthesis filter.

50. A method of decoding speech signal based on received feature parameters by changing a decoding operation according to whether the speech signal is in a voice period or in a voice-less period, at least a part of decoding process in the method in the voice-less period comprising the steps of:

smoothing the received feature parameters and calculating smoothed parameters;

determining a weighting coefficient used to generate an excitation signal of the voice-less period by performing a weighted sum operation of plural types of signals based on at least one of the

smoothed parameters; and

generating the excitation signal based on the weighting coefficient, and generating speech signal in the voice-less period by feeding the excitation signal to a synthesis filter.

51. The speech decoding device of claim 34, wherein the feature parameters includes at least one of a quantity representing spectral envelope of the signals to be decoded and a quantity representing power of the signals to be decoded.

52. The speech decoding device of claim 35, wherein the feature parameters includes at least one of a quantity representing spectral envelope of the signals to be decoded and a quantity representing power of the signals to be decoded.

53. The speech decoding device of claim 37, wherein the feature parameters includes at least one of a quantity representing spectral envelope of the signals to be decoded and a quantity representing power of the signals to be decoded.

54. The speech decoding device of claim 39, wherein the feature parameters includes at least one of a quantity representing spectral envelope of the signals to be decoded and a quantity representing power of the signals to be decoded.

55. The speech decoding device of claim 49, wherein the feature parameters includes at least one of a quantity representing spectral envelope of the signals to be decoded and a quantity representing power of the signals to be decoded.

56. The speech decoding device of claim 50, wherein the feature parameters includes at least one of a quantity representing spectral envelope of the signals to be decoded and a quantity representing power of the signals to be decoded.

57. A recording medium which records a program performing a method of decoding speech signal by changing a decoding operation corresponding to received feature parameters according to whether the speech signal is in a voice period or in a voice-less period, the method comprising the steps of:

smoothing the feature parameter representing spectral envelope characteristics of the speech signal to be decoded in at least a part of the voice-less period; and

decoding the speech signal in the voice-less period by using the smoothed feature parameter.

58. A recording medium which records a program performing a method of decoding speech signal by changing a decoding operation corresponding to plural types of received feature parameters according to whether the speech signal is in a voice period or in a voice-less period, the method comprising the steps of:

smoothing at least one of the feature parameters according to an elapsed time from a time point when a transition occurs from the voice period to the voice-less period; and

decoding the speech signal in the voice-less period by using the smoothed feature parameter.

59. The recording medium of claim 58, wherein the smoothing step further comprises the steps of:

(a) providing at least one of the received feature parameters as it is as the smoothed feature parameter in a certain time period immediately after changing from the voice period to the voice-less period; and

(b) smoothing at least one of the received feature parameters in a time period other than the certain time period.

60. A recording medium which records a program performing a method of decoding speech signal by changing a decoding operation corresponding to plural types of received feature parameters according to whether the speech signal is in a voice period or in a voice-less period, the method comprising the steps of:

smoothing at least one of the feature parameters according to the feature parameters; and

decoding the speech signal in the voice-less period by using the smoothed feature parameter.

61. The recording medium of claim 60, wherein the smoothing step further comprises the steps of:

(a) providing at least one of the received feature parameters as it is as the smoothed feature parameter while the feature parameter satisfies a predetermined condition; and

(b) smoothing at least one of the received feature parameters after the condition is not satisfied.

62. A recording medium which records a program performing a method of decoding speech signal by changing a decoding operation corresponding to plural types of received feature parameters according to whether the speech signal is in a voice period or in a voice-less period, the method comprising the steps of:

smoothing at least one of the feature parameters according to information representing whether the feature parameters are transmitted or not; and

decoding the speech signal in the voice-less period by using the smoothed feature parameter.

63. The recording medium of claim 58, wherein the smoothing step smoothes at least one of the received feature parameters according to an elapsed time from when a transition occurs from the

voice period to the voice-less period, and the feature parameters.

64. The recording medium of claim 58, wherein the smoothing step smoothes at least one of the received feature parameters according to an elapsed time from when a transition occurs from the voice period to the voice-less period, and at least one of the feature parameters, after at least one of the feature parameters is used as it is.

65. The recording medium of claim 60, wherein the smoothing step smoothes at least one of the received feature parameters according to an elapsed time from when a transition occurs from the voice period to the voice-less period, and at least one of the feature parameters, after at least one of the feature parameters is used as it is.

66. The recording medium of claim 58, wherein the smoothing step further comprises the steps of:

(a) providing at least one of the received feature parameters as it is, in a first time period immediately after changing from the voice period to the voice-less period and in a second time period while the feature parameter satisfies a predetermined condition; and

(b) smoothing at least one of the received feature parameters in a time direction after the first time period or the second time period is past.

67. The recording medium of claim 58, wherein the smoothing step changes a coefficient used to smooth at least one of the received feature parameters according to information representing whether the feature parameters are transmitted or not.

68. The recording medium of claim 60, wherein the smoothing step changes a coefficient used to smooth at least one of the received feature parameters according to information representing whether the feature parameters are transmitted or not.

69. The recording medium of claim 62 further comprising the step of receiving the information representing whether the feature parameters are transmitted or not.

70. The recording medium of claim 67 further comprising the step of receiving the information representing whether the feature parameters are transmitted or not.

71. The recording medium of claim 68 further comprising the step of receiving the information representing whether the feature parameters are transmitted or not.

72. A recording medium which records a program performing a method of decoding speech signal based on received feature parameters by changing a decoding operation according to whether the speech signal is in a voice period or in a voice-less period, at least a part of decoding process in the method in the voice-less period comprising the steps of:

determining a weighting coefficient used to generate an excitation signal of the voice-less period by performing a weighted sum operation of plural types of signals based on at least one of the received feature parameters; and

generating the excitation signal based on the weighting coefficient, and generating speech signal in the voice-less period by feeding the excitation signal to a synthesis filter.

73. A recording medium which records a program performing a method of decoding speech signal based on received feature parameters by changing a decoding operation according to whether the speech signal is in a voice period or in a voice-less period, at least a part of decoding process in the method in the voice-less period comprising the steps of:

smoothing the received feature parameters and calculating smoothed parameters;

determining a weighting coefficient used to generate an excitation signal of the voice-less period by performing a weighted sum operation of plural types of signals based on at least one of the smoothed parameters; and

generating the excitation signal based on the weighting coefficient, and generating speech signal in the voice-less period by feeding the excitation signal to a synthesis filter.